

Db	121	QKRYEQQREDEEKFERMEKGDNKRDPQREYECRRHCEQQERLQYQCQRQEQQR 180	Oy	121	QKRYEQQREDEEKFERMEKGDNKRDPQREYECRRHCEQQERLQYQCQRQEQQR 180
Qy	181	QHGRGDDLMNPORGSSGRYEGEEKQSDNYYFDRSLSTRFRTEEGHISVLENYGRSK 240	Db	162	QKRYEQQREDEEKFERMEKGDNKRDPQREYECRRHCEQQERLQYQCQRQEQQR 221
Db	181	QHGRGDDLMNPORGSSGRYEGEEKQSDNYYFDRSLSTRFRTEEGHISVLENYGRSK 240	Oy	181	QHGRGDDLMNPORGSSGRYEGEEKQSDNYYFDRSLSTRFRTEEGHISVLENYGRSK 240
Qy	241	LRLAKNYRVLLEANPNAFLPHTLDAILVIGGRGALKMTHRDNRHRSYNEECGDI 300	Db	222	QHGRGDDLMNPORGSSGRYEGEEKQSDNYYFDRSLSTRFRTEEGHISVLENYGRSK 281
Db	241	LRLAKNYRVLLEANPNAFLPHTLDAILVIGGRGALKMTHRDNRHRSYNEECGDI 300	Oy	241	LRLAKNYRVLLEANPNAFLPHTLDAILVIGGRGALKMTHRDNRHRSYNEECGDI 300
Qy	301	RIPAGITFYILNDRNNEERLHAKFLQTISPGQYKEFPAGQNEPEYLSTFSKILEA 360	Db	282	LRLAKNYRVLLEANPNAFLPHTLDAILVIGGRGALKMTHRDNRHRSYNEECGDI 341
Db	301	RIPAGITFYILNDRNNEERLHAKFLQTISPGQYKEFPAGQNEPEYLSTFSKILEA 360	Oy	301	RIPAGITFYILNDRNNEERLHAKFLQTISPGQYKEFPAGQNEPEYLSTFSKILEA 360
Qy	361	LNTQTERLRLGVLGQOREGVITRASOQEIRLTRODSESRWHRGESSRGPNLFNKR 420	Db	342	RIPAGITFYILNDRNNEERLHAKFLQTISPGQYKEFPAGQNEPEYLSTFSKILEA 401
Db	361	LNTQTERLRLGVLGQOREGVITRASOQEIRLTRODSESRWHRGESSRGPNLFNKR 420	Oy	361	LNTQTERLRLGVLGQOREGVITRASOQEIRLTRODSESRWHRGESSRGPNLFNKR 420
Qy	421	PLYSKNGQAYEVKEVKEPDYRQLQDMVSFLANITQGSMGPPFENRSTKVVVASGEADY 480	Db	402	LNTQTERLRLGVFGQOREGVITRASOQEIRLTRODSESRWHRGESSRGPNLFNKR 461
Db	421	PLYSKNGQAYEVKEVKEPDYRQLQDMVSFLANITQGSMGPPFENRSTKVVVASGEADY 480	Oy	421	PLYSKNGQAYEVKEVKEPDYRQLQDMVSFLANITQGSMGPPFENRSTKVVVASGEADY 480
Qy	481	EMACPHLSGRGGGRKRIHEEEEVHYPVFSSENLL 540	Db	462	PLYSKNGQAYEVKEVKEPDYRQLQDMVSFLANITQGSMGPPFENRSTKVVVASGEADY 521
Db	481	EMACPHLSGRGGGRKRIHEEEEVHYPVFSSENLL 540	Oy	481	EMACPHLSGRGGGRGCKRHEEEEVHYPVFSSENLL 540
Qy	541	LFAFGINAQNHHENFLAGRERNVLOQIEPOAMELAFAASRKEVEELFNSQDESIFFPGP 600	Db	522	EMACPHLSGRGGGRGCKRHEEEEVHYPVFSSENLL 581
Db	541	LFAFGINAQNHHENFLAGRERNVLOQIEPOAMELAFAASRKEVEELFNSQDESIFFPGP 600	Oy	541	LFAFGINAQNHHENFLAGRERNVLOQIEPOAMELAFAASRKEVEELFNSQDESIFFPGP 600
Qy	601	QHQQSPRSTKQQPLVSLIDFVGF 625	Db	582	LFAFGINAQNHHENFLAGRERNVLOQIEPOAMELAFAASRKEVEELFNSQDESIFFPGP 641
Db	601	QHQQSPRSTKQQPLVSLIDFVGF 625	Oy	601	QHQQSPRSTKQQPLVSLIDFVGF 625
Db	642	QHQQSPRSTKQQPLVSLIDFVGF 666	Db	642	QHQQSPRSTKQQPLVSLIDFVGF 666
RESULT	2				
Q9SPL5					
ID	Q9SPL5	PRELIMINARY; PRT; 666 AA.			
AC	Q9SPL5;				
DT	01-MAY-2000 (TREMBLrel. 13, Created)				
DT	01-MAY-2000 (TREMBLrel. 13, Last sequence update)				
DE	01-OCT-2000 (TREMBLrel. 15, Last annotation update)				
GN	VICILIN PRECURSOR.				
OS	Macadamia integrifolia (Macadamia nut)				
OC	Bukaryota; Viridiplantae; Embryophyta; Spermatophyta;				
OC	Magnoliophyta; eudicotyledons; Proteaceae; Macadamia.				
OX	NCBI_TaxID=60598;				
RN	SEQUENCE FROM N.A.				
RC	TISSUE=NUT KERNEL;				
RT	"A family of antimicrobial peptides is produced by processing of a 7S globulin protein in Macadamia integrifolia kernels."				
RL	Plant; J. 0-0 (1999); AAD4244.1; -.				
DR	EMBL; AF161883; AAD4244.1; -.				
DR	HSSP; P02833; 2PBL.				
DR	INTERPRO; IPR00113; -.				
PFAM; PF00546; Seedstore_7s; 1.					
NCBI_TaxID=60598;					
SEQUENCE FROM N.A.					
RC	TISSUE=NUT KERNEL;				
RT	"A family of antimicrobial peptides is produced by processing of a 7S globulin protein in Macadamia integrifolia."				
RL	Plant; J. 0-0 (1999);				
DR	EMBL; AF161884; AAD4245.1; -.				
DR	HSSP; P02833; 2PBL.				
DR	INTERPRO; IPR00113; -.				
PFAM; PF00546; Seedstore_7s; 1.					
NCBI_TaxID=60598;					
SEQUENCE FROM N.A.					
RC	TISSUE=NUT KERNEL;				
RT	"A family of antimicrobial peptides is produced by processing of a 7S globulin protein in Macadamia integrifolia."				
RL	Plant; J. 0-0 (1999);				
DR	EMBL; AF161884; AAD4245.1; -.				
DR	HSSP; P02833; 2PBL.				
DR	INTERPRO; IPR00113; -.				
PFAM; PF00546; Seedstore_7s; 1.					
NCBI_TaxID=60598;					
SEQUENCE FROM N.A.					
RC	TISSUE=NUT KERNEL;				
RT	"A family of antimicrobial peptides is produced by processing of a 7S globulin protein in Macadamia integrifolia."				
RL	Plant; J. 0-0 (1999);				
DR	EMBL; AF161884; AAD4245.1; -.				
DR	HSSP; P02833; 2PBL.				
DR	INTERPRO; IPR00113; -.				
PFAM; PF00546; Seedstore_7s; 1.					
NCBI_TaxID=60598;					
SEQUENCE FROM N.A.					
RC	TISSUE=NUT KERNEL;				
RT	"A family of antimicrobial peptides is produced by processing of a 7S globulin protein in Macadamia integrifolia."				
RL	Plant; J. 0-0 (1999);				
DR	EMBL; AF161884; AAD4245.1; -.				
DR	HSSP; P02833; 2PBL.				
DR	INTERPRO; IPR00113; -.				
PFAM; PF00546; Seedstore_7s; 1.					
NCBI_TaxID=60598;					
SEQUENCE FROM N.A.					
RC	TISSUE=NUT KERNEL;				
RT	"A family of antimicrobial peptides is produced by processing of a 7S globulin protein in Macadamia integrifolia."				
RL	Plant; J. 0-0 (1999);				
DR	EMBL; AF161884; AAD4245.1; -.				
DR	HSSP; P02833; 2PBL.				
DR	INTERPRO; IPR00113; -.				
PFAM; PF00546; Seedstore_7s; 1.					
NCBI_TaxID=60598;					
SEQUENCE FROM N.A.					
RC	TISSUE=NUT KERNEL;				
RT	"A family of antimicrobial peptides is produced by processing of a 7S globulin protein in Macadamia integrifolia."				
RL	Plant; J. 0-0 (1999);				
DR	EMBL; AF161884; AAD4245.1; -.				
DR	HSSP; P02833; 2PBL.				
DR	INTERPRO; IPR00113; -.				
PFAM; PF00546; Seedstore_7s; 1.					
NCBI_TaxID=60598;					
SEQUENCE FROM N.A.					
RC	TISSUE=NUT KERNEL;				
RT	"A family of antimicrobial peptides is produced by processing of a 7S globulin protein in Macadamia integrifolia."				
RL	Plant; J. 0-0 (1999);				
DR	EMBL; AF161884; AAD4245.1; -.				
DR	HSSP; P02833; 2PBL.				
DR	INTERPRO; IPR00113; -.				
PFAM; PF00546; Seedstore_7s; 1.					
NCBI_TaxID=60598;					
SEQUENCE FROM N.A.					
RC	TISSUE=NUT KERNEL;				
RT	"A family of antimicrobial peptides is produced by processing of a 7S globulin protein in Macadamia integrifolia."				
RL	Plant; J. 0-0 (1999);				
DR	EMBL; AF161884; AAD4245.1; -.				
DR	HSSP; P02833; 2PBL.				
DR	INTERPRO; IPR00113; -.				
PFAM; PF00546; Seedstore_7s; 1.					
NCBI_TaxID=60598;					
SEQUENCE FROM N.A.					
RC	TISSUE=NUT KERNEL;				
RT	"A family of antimicrobial peptides is produced by processing of a 7S globulin protein in Macadamia integrifolia."				
RL	Plant; J. 0-0 (1999);				
DR	EMBL; AF161884; AAD4245.1; -.				
DR	HSSP; P02833; 2PBL.				
DR	INTERPRO; IPR00113; -.				
PFAM; PF00546; Seedstore_7s; 1.					
NCBI_TaxID=60598;					
SEQUENCE FROM N.A.					
RC	TISSUE=NUT KERNEL;				
RT	"A family of antimicrobial peptides is produced by processing of a 7S globulin protein in Macadamia integrifolia."				
RL	Plant; J. 0-0 (1999);				
DR	EMBL; AF161884; AAD4245.1; -.				
DR	HSSP; P02833; 2PBL.				
DR	INTERPRO; IPR00113; -.				
PFAM; PF00546; Seedstore_7s; 1.					
NCBI_TaxID=60598;					
SEQUENCE FROM N.A.					
RC	TISSUE=NUT KERNEL;				
RT	"A family of antimicrobial peptides is produced by processing of a 7S globulin protein in Macadamia integrifolia."				
RL	Plant; J. 0-0 (1999);				
DR	EMBL; AF161884; AAD4245.1; -.				
DR	HSSP; P02833; 2PBL.				
DR	INTERPRO; IPR00113; -.				
PFAM; PF00546; Seedstore_7s; 1.					
NCBI_TaxID=60598;					
SEQUENCE FROM N.A.					
RC	TISSUE=NUT KERNEL;				
RT	"A family of antimicrobial peptides is produced by processing of a 7S globulin protein in Macadamia integrifolia."				
RL	Plant; J. 0-0 (1999);				
DR	EMBL; AF161884; AAD4245.1; -.				
DR	HSSP; P02833; 2PBL.				
DR	INTERPRO; IPR00113; -.				
PFAM; PF00546; Seedstore_7s; 1.					
NCBI_TaxID=60598;					
SEQUENCE FROM N.A.					
RC	TISSUE=NUT KERNEL;				
RT	"A family of antimicrobial peptides is produced by processing of a 7S globulin protein in Macadamia integrifolia."				
RL	Plant; J. 0-0 (1999);				
DR	EMBL; AF161884; AAD4245.1; -.				
DR	HSSP; P02833; 2PBL.				
DR	INTERPRO; IPR00113; -.				
PFAM; PF00546; Seedstore_7s; 1.					
NCBI_TaxID=60598;					
SEQUENCE FROM N.A.					
RC	TISSUE=NUT KERNEL;				
RT	"A family of antimicrobial peptides is produced by processing of a 7S globulin protein in Macadamia integrifolia."				
RL	Plant; J. 0-0 (1999);				
DR	EMBL; AF161884; AAD4245.1; -.				
DR	HSSP; P02833; 2PBL.				
DR	INTERPRO; IPR00113; -.				
PFAM; PF00546; Seedstore_7s; 1.					
NCBI_TaxID=60598;					
SEQUENCE FROM N.A.					
RC	TISSUE=NUT KERNEL;				
RT	"A family of antimicrobial peptides is produced by processing of a 7S globulin protein in Macadamia integrifolia."				
RL	Plant; J. 0-0 (1999);				
DR	EMBL; AF161884; AAD4245.1; -.				
DR	HSSP; P02833; 2PBL.				
DR	INTERPRO; IPR00113; -.				
PFAM; PF00546; Seedstore_7s; 1.					
NCBI_TaxID=60598;					
SEQUENCE FROM N.A.					
RC	TISSUE=NUT KERNEL;				
RT	"A family of antimicrobial peptides is produced by processing of a 7S globulin protein in Macadamia integrifolia."				
RL	Plant; J. 0-0 (1999);				
DR	EMBL; AF161884; AAD4245.1; -.				
DR	HSSP; P02833; 2PBL.				
DR	INTERPRO; IPR00113; -.				
PFAM; PF00546; Seedstore_7s; 1.					
NCBI_TaxID=60598;					
SEQUENCE FROM N.A.					
RC	TISSUE=NUT KERNEL;				
RT	"A family of antimicrobial peptides is produced by processing of a 7S globulin protein in Macadamia integrifolia."				
RL	Plant; J. 0-0 (1999);				
DR	EMBL; AF161884; AAD4245.1; -.				
DR	HSSP; P02833; 2PBL.				
DR	INTERPRO; IPR00113; -.				
PFAM; PF00546; Seedstore_7s; 1.					
NCBI_TaxID=60598;					
SEQUENCE FROM N.A.					
RC	TISSUE=NUT KERNEL;				
RT	"A family of antimicrobial peptides is produced by processing of a 7S globulin protein in Macadamia integrifolia."				
RL	Plant; J. 0-0 (1999);				
DR	EMBL; AF161884; AAD4245.1; -.				
DR	HSSP; P02833; 2PBL.				
DR	INTERPRO; IPR00113; -.				
PFAM; PF00546; Seedstore_7s; 1.					
NCBI_TaxID=60598;					
SEQUENCE FROM N.A.					
RC	TISSUE=NUT KERNEL;				
RT	"A family of antimicrobial peptides is produced by processing of a 7S globulin protein in Macadamia integrifolia."				
RL	Plant; J. 0-0 (1999);				
DR	EMBL; AF161884; AAD4245.1; -.				
DR	HSSP; P02833; 2PBL.				
DR	INTERPRO; IPR00113; -.				
PFAM; PF00546; Seedstore_7s; 1.					
NCBI_TaxID=60598;					
SEQUENCE FROM N.A.					
RC	TISSUE=NUT KERNEL;				
RT	"A family of antimicrobial peptides is produced by processing of a 7S globulin protein in Macadamia integrifolia."				
RL	Plant; J. 0-0 (1999);				
DR	EMBL; AF161884; AAD4245.1; -.				
DR	HSSP; P02833; 2PBL.				
DR	INTERPRO; IPR00113; -.				
PFAM; PF00546; Seedstore_7s; 1.					
NCBI_TaxID=60598;					
SEQUENCE FROM N.A.					
RC	TISSUE=NUT KERNEL;				
RT	"A family of antimicrobial peptides is produced by processing of a 7S globulin protein in Macadamia integrifolia."</td				

QY	61 CORCKEKECEEEENRQDPOQOYEQOCORRETEPRHMOCQCORCKEERKQ	120	QY	32 DNOEDPTECCOCQRCRQEQESDPRQOYCQRCRKECEE---EEENRQ-----	78
Db	102 CORCKEKECEEEENRQDPOQOYEQOCORRETEPRHMOTCQCKCERRYEKEKRQ	161	Db	9 ENERDPREQYRQGQCEQRQGQDQOCQRCRQEQEEDQRSQEERRGRDQDQ	68
Qy	121 QKREEQQREDEEEYERMEKGENKRDQDQREVEDCRHCEQEPRLQYQCORRCQQR	180	Qy	79 -RDQDQYEQOCRKCRRETEPRHMQLCQCRQRCRQYEEKRKQYEEQORREDEEYEE	137
Db	162 QKREEQQREDEEEYERMEKGENKRDQDQREVEDCRHCEQEPRLQYQCORRCQQR	221	Db	69 PRDPEQYEQCQOC-EQRGQEOTLCRRCEQRORQEEERERQR-----	112
Qy	181 QHGRRGGDLMNPORGSGRVEEGEEKQSDNPYFDERSI-STRFREEGHSVLENFYRSK	240	Qy	138 RAKEGDANKRDQDQREVEDCRHIC--EQEPRLQYQCORRCQ-QRQHGR-GGDLMNQR	193
Db	222 QHGRRGGDLMNPORGSGRVEEGEEKQSDNPYFDERSI-STRFREEGHSVLENFYRSK	281	Db	113 ---GRDQDPOQ-QYHCRQRCQIQOSPERQRQCORCERQYBQGRRGPEASPR	167
Qy	241 LLRALKNTRVLVLEAANPVAFLPDTLDAATLVLIGRGLAKMHRNRESYNECQDVI	300	Qy	194 GGSGRYEEGEKQSDNPYFDERSI-STRFREEGHSVLENFYRSK	253
Db	282 LLRALKNTRVLVLEAANPVAFLPDTLDAATLVLIGRGLAKMHRNRESYNECQDVI	341	Db	168 ESGR--EEQCRHNPYFQHISQISRHSRHESEGEVYLERTERTELLRGENYRVL	224
Qy	301 RIGATTYLINEDNEELHAKFLQITSPGQKEFPAGGONPEYLTSFSKELLEA	360	Qy	254 EANPNAFLPHTLDAATLVLIGRGLAKMHRNRESYNECQDVI	313
Db	342 RIGATTYLINEDNEELHAKFLQITSPGQKEFPAGGONPEYLTSFSKELLEA	401	Db	225 DANPNTSMPLPHKDAESWAVIWRGRATLTVLSQETRESFNLCGDIVRVPAGATVING	284
Qy	361 LNTQTERLRLVGLGQREGVITRASSEQIRELTTRDSESRRWHLTRGGBESSRGPYLNKR	420	Qy	314 DANNERLHAKFLQITSPGQKEFPAGGONPEYLTSFSKELLEA	372
Db	402 LNTQERLRLVGLGQREGVISASEQIRELTTRDSESRRWHLTRGGBESSRGPYLNKR	461	Db	285 DSNERLENKVKLQVNPUNIPQFQFEEYAAKPSDQSYLIVFNDLVAALNTPDRJEE	344
Qy	421 PLXNSKYQAYEVKPEDYRQLQDMDSVFTANITQGSMGPFNTSTKVVVAVGADY	480	Qy	373 GQ---OREGVITRASSEQIRELTTRD---DSESRWHLTRGGESSRGPYLNPLNPRLYSNKG	428
Db	462 PLXNSKYQAYEVKPEDYRQLQDMDSVFTANITQGSMGPFNTSTKVVVAVGADY	521	Db	345 DQEORECVITRASSEQIRELTTRDSESRRWHLTRGGESSRGPYLNPLNPRLYSNKG	400
Qy	481 EMACPHLSGRHGRRGGKRRHEEEVYEVQVARRLSREATVLAGIPWVFSGNELL	540	Qy	429 QAEVKPEDYRQLQDMDSVFTANITQGSMGPFNTSTKVVVAVGADY	488
Db	522 EMACPHLSGRHGRRGGKRRHEEEVYEVQVARRLSREATVLAGIPWVFSGNELL	581	Db	401 QFFEACPFEEHRLQMDVLYNVAEIKRCAMMVFPHNSKATVVVVBGTGRYEMACPHVS	460
Qy	541 LFFPGINAQNNHENFLAGRERNVLLQTEPQAMELAFAASRKEVEELENSQDESIFFGPGR	600	Qy	489 GRIGGGGGKGRHEEEEVY---HVEQVARRLSREATVLAGIPWVFSGNELLFAFGI	546
Db	582 LFFPGINAQNNHENFLAGRERNVLLQTEPQAMELAFAASRKEVEELENSQDESIFFGPGR	641	Db	461 SQ-SYEQGRQDQEELSTGRQVITARLQDIFVTPAGHLATIASQNENLRLGFDI	519
Qy	601 OHQOQSRSRSTKOOQPLVSILDFVGF	625	Qy	547 NAQNNHENFLAGRERNVLLQTEPQAMELAFAASRKEVEELENSQDESIFFGPGR	606
Db	642 OHQOQSRSRSTKOOQPLVSILDFVGF	666	Db	520 NGENNQDFLAG-QNNTINOLEREAKELSFNMPRELEEEESQMSFYVPTER---QS	574
RESULT 4					
OPSEW4	PRELIMINARY;	PRT;	593 AA.		
ID	OPSEW4				
AC					
DT	01-MAY-2000 (TREMBLRL. 13, Created)				
DT	01-MAY-2000 (TREMBLRL. 13, Last sequence update)				
DT	01-OCT-2000 (TREMBLRL. 15, Last annotation update)				
DE	VICILIN-LIKE PROTEIN PRECURSOR (FRAGMENT).				
OS	Juglans regia (English walnut).				
OC					
OX	NBI_MaxID=51240;				
RN	[1]				
RP	SEQUENCE FROM N-A.				
RC	STRAIN-CV: SUNLAND; TISSUE=SOMATIC EMBRYO LINE;				
RA	Teuber S.S., Peterson W.R., Dandekar A.M., Ansari A.A.;				
RT	"Identification and cloning of a cDNA encoding a vicilin-like protein, Jug r 2, from English walnut kernel (Juglans regia): a major food allergen," (NAY-1998) to the EMBL/GenBank/DDBJ databases.				
RT	Submitted: AF066055; RAf18269; 1; .				
RT	HSSP: P02853; 2PML.				
RT	INTERPRO: IPR001113; -.				
RT	PFAM: PF00546; Seedstore_7s; 1.				
RT	NON_TER_1				
RT	SEQUENCE FROM N-A.				
RL	STRAIN-KUROKAWA: AMAKURI NANKIN; TISSUE=COTYLEDON;				
DR	EMBL: AB019195; RA34056; 1; -.				
DR	HSSP: P02853; 2PML.				
DR	INTERPRO: IPR001113; -.				
DR	PFAM: PF00546; Seedstore_7s; 1.				
DR	PRODOM: PD08105; -; 1.				
DR	SEQUENCE 593 AA; 69990 MW; 9BA127E19B18C0D8 CRC64;				
SQ	Query Match 39.4%; Score 1310.5; DB 10; Length 593; Best Local Similarity 44.1%; Pred. No. 4.8e-88; Matches 273; Conservative 121; Mismatches 166; Indels 59; Gaps 17;				

Query Match	35.0%	Score 1163; DB 10; Length 810;	RP	SEQUENCE FROM N.A.
Best Local Similarity	36.8%	Pred. No. 4 9e-7;	RC	TISSUE-LEAVES;
Matches	271;	Conservative 133; Mismatches 192; Indels 140; Gaps 22;	RX	MEDLINE=92288309; PubMed=1600151;
Oy	10	QMRVRSQCDERFEE-----DIDNARY----- 31	RA	"Comparison of the structure and nucleotide sequences of vicilin genes of cocoa and cotton raise questions about vicilin evolution.";
Db	94	QORKCQEVCERLREREQRGEDVDEVERDPEWEREQRREHREERRRERER 153	RT	Plant Mol. Biol. 18:117-1176 (1992).
Oy	32	-----DNEQEPQTECQQCQRCRQESDPQQYQQRCKEICEEEVNRQRPQQY 85	RL	McHenry L., Fritz P.J.;
Db	154	GRGRDRENERIPKREOEROR--REQERRREQERERERRGERBEDDEN-QRDPDWRR 210	DR	EMBL: X62625; CAA44494.1; -.
Oy	86	EQCQKRCQRETEPRIMQICOERCERY-----EKE-----KRKQQKRYEQQREDE 132	DR	EMBL: X62626; CAA44494.1; -.
Db	211	EQ----ERQEQRAREQEQCERQRGGRDDDENQRDPDWREQKREQEQCER 265	DR	INNSP; P02853; 2PBL.4
Oy	133	ERYERKREGONKRD-----QREYEDCRRHCTQ----- 166	DR	MENDELE, 30919; Thecc:1188-30919.
Db	266	ERRGRGRDDEDENQRDPDWRRQERREQERRREREHRSGRDVEENQRPDW 325	DR	INTERPRO; IPR001113; -.
Oy	167	LOYOCORRCOPO-----QROHGRG----- 166	PFAM; PF00546; Seedstore_7s; 1.	
Db	326	REQERRKREEQREWEREHRGRRGREGQRSREDERRHEROGGRSRVNVQAIRTEQE 385	DR	PRODOM; PDO81059; -; 1.
Oy	206	QSDNPYFDELSLSTRTECHISYLENPFGRSKLRLKNYRVLVLEANPNAVLPHT 265	KW	SIGNAL.
Db	386	QSNNPYFOEGRFOFSYRSDEGHWRVLERSERSELLKGKPHNORVALLARPHTFVPH 445	FT	1
Oy	266	LJADADLLVIGRAKMIHRD-----ESTNLCODVIRIPAGTTFYLINRDNNERBLH 321	CHAIN	24
Db	446	LDAECVLLVGRGATITVWQEKFRKESSTINVESSDVMITPAGTFLYLANQE NEDQI 504	SEQUENCE	525 AA; 60798 MW; 19114CD5C248905D CRC64;
Oy	322	AKFLORTISTPOQYKEPPASQNPPEYLSTPSKELAANTOTERLURGLQOQE-----GV 379	Query Match	30.5%; Score 2015; DB 10; Length 525;
Db	505	VKLVQPVNNPGEFKYDLSAGGES-QAYVSVSNDVLAELALNIPRDKLERIFKORKRGGK 563	Best Local Similarity	41.1%; Pred. No. 1 9e-66;
Oy	380	ITRASQEQLRTRDOSSESRWHRIGGESSRGPGYMLNKRPLYNSKQGAYPEVVEDYR 439	Matches	211; Conservative 91; Mismatches 158; Indels 54; Gaps 11;
Db	564	IVRASOBLRALSQ-----RATSVRKGSRGVAPRKLESQTPVYNNQYGMFEACDEFP 618	Qy	68 ICEEEEYNR--QRPQQQYEQOKRCORRETEREPHMOCQQRCCERRYSEKKRQQKRY 124
Oy	440	OLQDMVSVFTANITGSMQCPFFNTRSTKVWVFAEGGSEFEMAPHIQSQWQRGEE 498	Db	22 ICSGVSYAGRQYERDPQQYEQCORCESATESEQEQQEQRER-----EV 70
Db	619	QLRRTQVATSVWDIKOGGMMPHFNRSRATWVVFVSEBAGSEFEMAPHIQSQWQRGEE 678	Qy	125 ERQOREDEEKFERMKEGDNRDPQOQEYEDCRRHCEQE-----PRQYOCORCQOQROH 182
Oy	499	---RUEEEE----VHYEQARLISKREAIWVLAGHPVVFVSSGENNLIFARGINAQ 549	Db	71 KEQRQEEE-----QROQOCQRCQEQOGQORQOCQRCQWQYKEQ 116
Db	679	ERHWRREEEEERERGRFERYAGRLUSEGGVLVLPAGHPTAIMASPNENLRLVGGINA 738	Qy	183 GKGGLDMLPQKQGSGHVEEGEKKQSDNPYDE-RLSLTRTERECHISYLENPFGRSKL 241
Oy	550	NNHENTLLAGRNRVNLQOIEPAMELAFAASKEVEELFNSODESFFPGPQRHQDQSPRS 609	Db	117 ER-GERENHYNHHKNRNSSEEGQQRNPNYFPKRSQFQTRDEGECKNFKILQRFIENSP 175
Db	739	NNHNRFLAGRE-NIMNEDRAKELAJAFNVEQKQADBFIRSQRESFTEGP----BGRRR 793	Qy	242 LRAKKNYRLVLEANPNAFVLPHTLADATLWVIGRGLAKMIIHRDNRNSYNELEGDVIR 301
Oy	610	TKOQPLSILDFVG 625	Db	176 LKGINDYRLAMFEANPNTFILPHICDAEATYFVTNGKTTFVTHENKESINVQRTVVS 235
Db	794	STERPLSLIILKLAGY 809	Qy	302 IPAGTTFYLINRDNNERBLHAKFLQTISTPQYKEPPAGCQNPPEYLSSKKELEAL 361
RESULT	6		Db	236 VPAGSTWVVSQDNOQKLTTAVLALPVVNSPKYELFPGNPKNPSSYGAFFSYEVLETF 295
ID	043358	PRELIMINARY; PRT; 525 AA.	Qy	362 NTQTERLRYGVQQR-----EGVIRASQEQLRELTRDOSSERRWHRIGGESSSRGP 413
AC	043358;		Db	296 WQREKLELIEQORQKROQGQGMFRKAKPQEQRQASQATSR----HRGGE-RLA 349
DT	01-NOV-1996 (TREMBLrel. 01, Last sequence update)		Qy	414 YNLFNKRPLYNSKQGAYPEVVEDYRQLODMDVSYIANITQGSMGMPFNTRSTKVVV 473
DT	01-OCT-2000 (TREMBLrel. 15, Last annotation update)		Db	350 INLLSOSPVSQNQGRFEACPEDSFQOFQNDMVAASFKLQNGAIFVPHYNISKATFVVF 409
DE	VICILIN PRECURSOR.		Qy	474 ASGEADVEMACPHLSRRGGGGK-----RHEEEV-----HYEQARLISKREIWL 525
GN	CSV.		Db	410 TDGYGYAQMACPHLSRQSOSSGRODRGEESEEETFGFEDQVKAPLSPGDVFWAPA 469
OS	Theobroma cacao (Cacao).		Qy	526 GIPVVFVSSGMLLFAFGNINAONNHNENFAAGR 559
OC	Bukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosid II; Malvales; Malvaceae; Theobroma.		Db	470 GHAVTFFASKQDPLNAVAFGJNAQNQRIEAGR 503
OC	NCBI_TaxID=3641; Theobroma.			
OC	zea mays (Maize); Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;			
OC	zeta mays (Maize); Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;			

OC Magnoliophyta; Liliopsida; Poales; Poaceae; Zea.
 NCBI_TaxID=4577;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=W64A;
 MEDLINE=92090707; PubMed=1752424;
 RA Belanger F.C.; Kriz A.L.;
 RT "Molecular basis for allelic polymorphism of the maize Globulin-1
 gene.";
 Genetics 129:863-872(1991).
 EMBL: X59082; CAA41809.1; -.
 HSSP: P50477; 1CNU.
 MENDEL: 11234; Zeama;1188;11234.
 DR INTERPRO; IPR000901; -.
 DR PROSITE; PS00546; Seedstore_7s; 1.
 DR PROSITE; PS00867; CPSASE_2; UNKNOWN_1.
 DR PRODOM; P0081059; -; 1.
 KW Seed storage protein.
 SEQUENCE 582 AA; 66162 MW; 680DB5FEC07CB885 CRC64;
 SQ

Query Match 26.7%; Score 888; DB 10; Length 582;
 Best Local Similarity 36.7%; Pred. No. 4; Se-57;
 Matches 208; Conservative 87; Mismatches 181; Indels 90; Gaps 13;
 QY 73 EFYRNQRDPQOQYEQCOKCORRETEPRHMO-ICQCORCERYEKRKOKORYE-SCORED 131
 Db 25 EDDNNHHHHGGHKSGQCVRRCEDR--PWHQRPRCLEQ-C-REERERKQERSRHEADRSQ 80
 QY 132 EBYKEYERMKEGDNKRPDQOQYEQCOKCORRETEPRHMO-ICQCORCERYEKRKOKORYE-SCORED 191
 Db 81 EGSSED------EREQE----- 91
 QY 192 ORGGSGSRYEEGEKQSD-NPYFEDRSLSLSTRFTEEGHISVLENYGRSKLRAKNYRL 250
 Db 92 -----REKQKDRRPFYFDRRSFRRVRSFEGSLRVLRRPDEVSRLRGIRDYRV 140
 QY 251 VILLEANPNAFVLPHTHADATLVLGGRGALKMHRDNRESYNYLQEGDVTRIAGTFYL 310
 Db 141 AVLEAANPSEFVWPSHDACTCYVAEGVWTTIENGERRSYTIQGHVFVAPAGAVTYL 200
 QY 311 INRNDNERHLAKEFQIYSTGQYKREPPAGQNTPEPYLSTPSKELAALNTOPRLRG 370
 Db 201 ANTGDGRKKLVTKLHTISVPGFQFFGGPGGRNPFSLFSFSKSTORAKYKISSDRLER 260
 QY 371 VLGO--OREGYTIRASQEQLRELTQDSE--SRWHIRGGESSRGPYNLFNKRPLYSN 425
 Db 261 LFGRHGDKGIGLIVRATTEQTRELRHRASEGQHGPWMLPPLPFGE-SRGPSILQDQPSIAN 319
 QY 426 KYQOAYEVKEPKDYLQLODMDSVFEANTICOSMMGPFNFNTSKVVAASGEADEVEMACP 485
 QY 320 QHGOQXLEADARSFHVLDVSYSPFANITAGSMSAPLNTRFSKFTAYVPGNGKVAEIVC 379
 QY 486 HLSGRG--GRGGKRRHFEEH-----YEQPARLRSKREATVVLAGHPVW 531
 Db 380 HRQSGQGESERERDGKRSBEEFESSEGEAEGQGYHTIRALSPGTAEVWPGHPEVA 439
 QY 532 VSSGNINLLFAGFGINAQNHNENPFLAGRERNVLOQIEPQAMELAASRSRKEVEELFNSOD 591
 Db 440 VNSRSDNLOIWCFFEVHADRKEVFLAGAD-NVQLDRVAKALPSASKABEVDEVLGSR 498
 QY 592 ESIIFPGPRO---HOQOSPRSTKQQ 614
 Db 499 EKGFLPGPPESESGGHEREQEDEEREE 524
 RESULT 8
 003678 PRELIMINARY; PRT; 637 AA.
 ID 003678
 AC 003678;
 DT 01-NOV-1996 (TremblieL 01, Created)
 DT 01-NOV-1996 (TremblieL 01, Last sequence update)

Query Match 26.0%; Score 864.5; DB 10; Length 637;
 Best Local Similarity 35.0%; Pred. No. 2.7e-55;
 Matches 221; Conservative 100; Mismatches 219; Indels 91; Gaps 19;
 QY 29 SKYDNOEDPQ-TECQCORRCRQQESEPSDPRQQYCQRRCKEICEBEEYNRQDQQYE 86
 Db 27 ASHDDDEDRGGHSILOQCVCQRQER-PR--YSHARCVQEC-----RDQDQH- 71
 QY 87 OCQKRCORRETEPRHMO-ICQCORCERYEKRKOKORYE-SCORED 146
 Db 72 -----GRHEQEEGGRGKRWGHEGEREEHGRGRHGEGEREEHGRGRH 119
 QY 147 DPOQREYEDCRRHRCOQEPRQYDQCORREOEOQORHGRG--GDLNIPQRGGSGRVEEG 202
 Db 120 GEGEREERGGHGGRGE-----REERERGRHGGEGEREEERGRGRGREG 170
 QY 203 E-EKQSDN--PYFEDRSLSLSTRFTEEGHISVLENYGRSKLRAKNYRLVILEANP 258
 Db 171 ERDEBQGDSRSPRYPVGPSSRFRITIQSDFHGVFRAIRPFDQYSLRIGIRDYRVIAEVNPR 230
 QY 259 AFVLPTHADAILVLGGRGALKMHRDNRESYNYLQEGDVTRIAGTFYLINRDNNER 318
 Db 231 AFVWVGFTDADGVGVQAQGGVLTENGERRSYTIQGVKEDVIVAPAGSILHNLANTDGRK 290
 QY 319 LHTAKFLOTISTPGQYKEFPPAGGONPEPPLSTSKEIIEALNTQTELRGVQGQ-- 375
 Db 291 LVIAKLHTISVPGKE-QFL---SVKPLIASLSSKVRRAFKTSDELERLQNQRCQ 344
 QY 376 --REGVITRASQEIRELTDDE--SRRWHIRGGESSRGPYNLFNKRPLYSNKGQ 429
 Db 345 EKTRSVSIVRASEEOLREIRREAEGQGQHWRPLPFRGDSRDTFNLBQRPKIANHGR 404
 QY 430 AYEVREDYFQLQDMDSVFEANTICOSMMGPFNFNTSKVVAASGEADEVEMACPHL-- 487
 Db 405 LYEADARSFLHANQDVYAWANTIPGSMFTAINTQSKLAVLLEGEGVQIVCPHLGR 464
 QY 488 --SGRHGG-----GGKRHPEEVHVEQYARLISKREAVVLAG 526
 Db 465 ESESERHGGKRRBEEEDDQQRGRGSESSEEEGQYARVRSRGSAPVPG 524
 QY 527 HPWVFVSS--GENENLILFAGFGINAQNHNENPFLAGRERNVLOQIEPQAMELAASRSRKEV 584
 Db 525 HPWVLEISSLSSQGSNSNLQVQVCFEINARNERNVWLAGR-NNVIGKIGSPAOELTFRGPAREVO 583
 QY 585 ELFNSQDESI-FFPSPRQHQSRSTKQQ 614

Db	584 EVFRAQDQDEGFVAGP---EQQSREQEQQE 611	Qy	587 FNSQDESIFPPGPRHQQQSPRSTKQQPLVSL 620
RESULT	9	;	;
ID	Q9M3X6	PRELIMINARY;	PRT; 613 AA.
AC	Q9M3X6;		
DT	01-OCT-2000 (T-EMBL; 15, Created)	022120	022120 PRELIMINARY; PRT; 544 AA.
DT	01-OCT-2000 (T-EMBL; 15, Last sequence update)	022120;	
DE	CONVICILIN PRECURSOR.	AC	022120;
GN	CVC.	DT	01-JAN-1998 (T-EMBL; 05, Last sequence update)
OS	pisum sativum (Garden pea).	DT	01-OCT-2000 (T-EMBL; 15, Last sequence update)
OC	Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids I;	DE	ALPHA SUBUNIT OF BETA CONGLYCININ.
OC	Fabales; Fabaceae; Papilionoideae; Pisum.	OS	Glycine max (Soybean).
OC	NCBI_TaxID=388;	DR	Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids I;
RN	[1]	OC	Fabales; Fabaceae; Papilionoideae; Glycine.
RP	SEQUENCE FROM N.A.	DR	NCBI_TaxID=3847;
RC	STRAIN=CV; TISSUE=SEED COTYLEDON;	RP	[1]
RA	Casey R.;	RC	SEQUENCE FROM N.A.
RT	"the sequence of a pea convicilin cDNA."	RA	STRAIN=CV; WASISUZUNARI;
RL	Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.	RL	Maruyama N., Katsube T., Wada Y., Utsumi S.;
DR	EMBL: AJ216875; CAB2855; 1; -.	DR	Submitted (NOV-1997) to the EMBL/GenBank/DBJ databases.
KW	SIGNAL.	DR	EMBL: AB008678; BAA23360; 1; -.
FT	SIGNAL	DR	HSSP; P50477; 1CAU.
FT	CHAIN	DR	MENDEL; 25074; Glyma; 1188; 25074.
FT	SEQUENCE	DR	INTERPRO; IPR001113; -.
SO	613 AA; 72063 MW; 49AED99F6135DD19 CRC64;	DR	PFAM; PF00546; Seedstore_7s; 1.
Query Match	25.4%; Score 844.5; DB 10; Length 613;	DR	PRODOM; P00810; 9; -; 1.
Best Local Similarity	32.6%; Pred. No. 7.5e-54;	DR	SEQUENCE 544 AA; 63296 MW; 04BB243929A9C9AB CRC64;
Matches	207; Conservative 122; Mismatches 216; Indels 89; Gaps 15;	SO	
Qy	26 IDWSKYNDNQDPOTPCQCORRCQESPRQOQYCORRKEICEEEEENRQ----- 79	Query Match	24.9%; Score 827; DB 10; Length 544;
Db	26 VTYANLYDEGSEPRVQAQERERGKEEERHGEWHRPSYE- EDEEGQERGRQE 83	Best Local Similarity	32.2%; Pred. No. 1.2e-52;
Qy	80 -----DQQQYFOCQK RCQRETEREPHMQICQQRCCERRYEKRKQRYE 125	Matches	187; Conservative 133; Mismatches 192; Indels 68; Gaps 12;
Db	84 KEEKRHGEWHRPSYEQKEDEEKQKRYOREKEDEEKQKRYOREKKEQEVQGRERWE 143	Qy	66 KEICEEEE--EYNQRDPQOCQYEQCQKRCQRCRRETEREPHMQICQQRCCERRYEKRKQK 122
Qy	126 EQQRDEEKHFERMKGNDNPFQREYEDCRRHICQEQQEPLQOCQRCQEQQRH 185	Db	4 KEECEBGETIPRPRPRPQHPRPQEPQGEKEBEDEDEQPRPPIPFPQPRPQPROPEEEHQREBQE 63
Db	144 RE--DEEQDDEEWGQSQRKEDPEE-----ARLHHEER -TKRRHH-- 183	Qy	123 RYEQOREDEEKEYERMKEDCNDKRPQDQREYEDCRRHICQEQQEPLQYQOCQRCQEQQRH 182
Qy	186 GDLMNPORGSGSGRYEGEEQKOSDNYYFDERSLSLSTRFRTEGHISVLENFGRSKLLRAL 245	Db	64 WPKKEEKGKGEKSEBEDEDEQEDERQPP-- RPHQHKEKQEEDED-- EEQRES 118
Db	184 -----QRESEBEERSESQERRNPFPLFSKNKFLFENENGHFLRQEDFKRSDFENL 237	Qy	183 GRRGDLMNPORGSGSGRYEEGEEKQ-----SDNPVYFDERSLSLSTRFRTERGHISVLENFG 237
Qy	246 KNYRLVLEUNPNAVYUPTLDADALILYIGGRALKMHRDNHESYNEFCGYTRIPAG 305	Db	119 -----ESEEDSELURHKNKNPPLFGSNRFETLEFKNQYGRIRQLRFQ 161
Db	238 QNYRLVYERAKPHTFLPQHIDALILWVLUSKGKLTWSPNDRNSYNLERGDTKLPG 297	Qy	238 RSKLLRALKNRYVLVLEUNPNAVYUPTLDADALILYIGGRALKMHRDNHESYNEFCGYTRIPAG 297
Qy	306 TTYKVLINRCONNERLHAKFQQTISTPGQKEFFAGQGPEPISTFSQEILEAMNQT 365	Db	162 RSPQQLNQLRDRYRILEFSKENTLILPNNHADADYLVLVINGTATLVLVNNDRDSYRLQSG 221
Db	298 TTSVLYNQDDEEDLILWLVIPNGPGKEAEDLA -KKNKNQYERGSKNILEYNTY 355	Qy	298 DVIRPAGTIFYLINRDNRNLHAKFQQTISTPGQKEFFAGQGPEPISTFSKEL 357
Qy	366 ERLRQVLGQRE-----GVITRASQEQTRELTRDSESRRWHIRRGGESSRG 413	Db	222 RSKLLRALKNRYVLVLEUNPNAVYUPTLDADALILYIGGRALKMHRDNHESYNEFCGYTRIPAG 281
Db	356 ETIWKVLEQEKRRKROGEETDAIKVSQEJELKLAKSS --KKSLSPEF 411	Qy	358 EAALNTQTERLRGV-----GOQR --EGVITRASQEQTRELTRDSESRRWHIRRGG 406
Qy	414 YNLFLNKRPIPSYNSKQYQAYEVKE-DYRQDMDYVFIANITOCQSMGPFNFTRSTKVV 472	Db	282 EASYDKFEELINKLFLSREBQQGQGEQRLQESVLEISEQIRSLSKRASSRKTI-- 338
Db	472 VNEKGKGNLEL-----GLKNEQERERKERNEVQYEARLSPGDWVTPG 519	Qy	407 GESSRGPYNLNKRPLPSYNSKQYQAYEVKE-DYRQDMDYVFIANITOCQSMGPFNFTRSTKVV 466
Db	412 INLRSHKPEYSNKFKLFEITPEKKYPOQLDQFLDVFSCVEINEGALMPLHYNSRAIVL 471	Db	339 -SSEDQPKPNLRSRDPYTSNKLQKPFPEITPEKPNPQLRDLDFLTSVDMNBDGALLPHENK 397
Qy	473 VASSEADVNACPHLSGRGGGGKRRHEEEVHIE---QVARAKSKREAVVLAG 526	Qy	467 STKVVVASGEADEMACPHLSGRGGGGKRRHEEEVHIE---VVARLSKREA 520
Db		Db	398 AIVLIVNEEDDANELV-----GLKNEQERERKERNEVQYEARLSPGDWVTPG 444
Qy	527 HPVVFVSSGENENLILFAFGINAONNHNENLAGRERVQIOPAMELAFAASRKEVEEL 586	Qy	521 IVVLAGHPVVFVSSGENENLILFAFGINAONNHNENLAGRERVQIOPAMELAFAASR 580
Db	520 HPVATASSNLNL--GFGINAENNERNFLSGDDINVSIQENPVWELTFPGSVQEINRL 577	Db	445 FVIRAGYIPVV--NATSNLNFAGINAENNRNFLAGSDQNNISOPSQWELAFCGA 502

Search completed: March 1, 2001, 16:09:23

